

University of Groningen

## Amorpha-4,11-diene synthase

Wallaart, TE; Bouwmeester, HJ; Hille, J; Poppinga, L; Maijers, NCA; Bouwmeester, Harro J.; Maijers, Niels C.A.

*Published in:*  
Planta

*DOI:*  
[10.1007/s004250000428](https://doi.org/10.1007/s004250000428)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2001

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Wallaart, TE., Bouwmeester, HJ., Hille, J., Poppinga, L., Maijers, NCA., Bouwmeester, H. J., & Maijers, N. C. A. (2001). Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. *Planta*, 212(3), 460-465.  
<https://doi.org/10.1007/s004250000428>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin

T. Eelco Wallaart<sup>1</sup>, Harro J. Bouwmeester<sup>2</sup>, Jacques Hille<sup>3</sup>, Lucas Poppinga<sup>1</sup>, Niels C. A. Majiers<sup>1</sup>

<sup>1</sup>GenoClipp biotechnology B.V., Meditech Center, L.J. Zielstraweg 1, 9713 GX Groningen, The Netherlands

<sup>2</sup>Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

<sup>3</sup>Department Molecular Biology of Plants, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received: 8 June 2000 / Accepted: 21 August 2000

**Abstract.** The sesquiterpenoid artemisinin, isolated from the plant *Artemisia annua* L., and its semi-synthetic derivatives are a new and very effective group of antimalarial drugs. A branch point in the biosynthesis of this compound is the cyclisation of the ubiquitous precursor farnesyl diphosphate into the first specific precursor of artemisinin, namely amorpha-4,11-diene. Here we describe the isolation of a cDNA clone encoding amorpha-4,11-diene synthase. The deduced amino acid sequence exhibits the highest identity (50%) with a putative sesquiterpene cyclase of *A. annua*. When expressed in *Escherichia coli*, the recombinant enzyme catalyses the formation of amorpha-4,11-diene from farnesyl diphosphate. Introduction of the gene into tobacco (*Nicotiana tabacum* L.) resulted in the expression of an active enzyme and the accumulation of amorpha-4,11-diene ranging from 0.2 to 1.7 ng per g fresh weight.

**Key words:** *Artemisia* – Artemisinin – Antimalarial drug – Amorpha-4,11-diene synthase

## Introduction

Artemisinin, a new and very potent antimalarial drug, is produced by the plant *Artemisia annua* L. in relatively small amounts ranging from 0.01% to 0.5% on a dry-weight basis (Wallaart et al. 2000). This makes artemisinin an expensive drug. Several studies describe the chemical synthesis of this sesquiterpene, but none of

these are an economical alternative for isolation of artemisinin from the plant (Webster and Lehnert 1994; Van Geldre et al. 1997). Therefore, a higher artemisinin concentration in the plant is desirable to make artemisinin available as a relatively cheap antimalaria drug. Knowledge of the biosynthetic pathway of this sesquiterpene may enable us to influence its formation in a direct way, for example by metabolic pathway engineering.

It has been postulated that the regulatory step in the biosynthesis of sesquiterpenes is represented by the cyclisation of the ubiquitous precursor farnesyl diphosphate (FDP) to the highly specific olefinic sesquiterpene skeletons (McCarvey and Croteau 1995). Furthermore, terpene synthases, like most enzymes involved in the biosynthesis of secondary metabolites, are known to occur in very low intracellular concentrations and to catalyse reactions rather slowly (Croteau and Cane 1985). The accumulation of artemisinic and dihydroartemisinic acids (Fig. 1) in the plant, and the absence of any intermediates en route from FDP to these two compounds, indicates that the terpene synthase-catalysed step is indeed rate limiting and may represent a regulatory point. Hence, cloning of the gene encoding amorpha-4,11-diene synthase and overexpression of this gene in *A. annua* seems to be the most rational approach to elevate the artemisinin level. To prove the functional expression of recombinant amorpha-4,11-diene synthase in a plant, a non-amorphadiene-producing plant should preferably be used for transformation. Tobacco is free of any amorpha-4,11-diene synthase and, because of its good transformation characteristics, is also a good candidate.

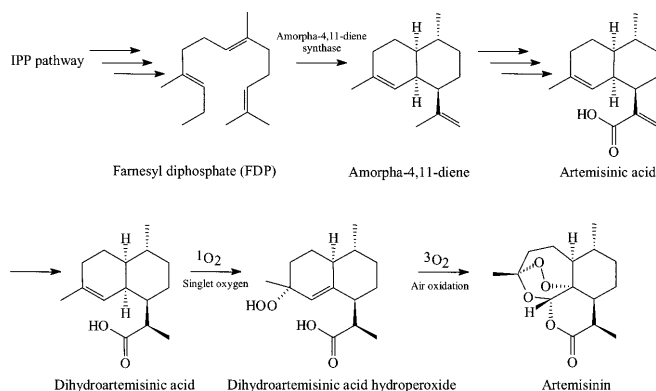
The cloning of terpene synthases can be achieved using a method based on the homology within these enzymes (Colby et al. 1993). The polymerase chain reaction (PCR) primers can be designed and used in a PCR with cDNA of *A. annua* as a template. The PCR product can subsequently be used as a probe for screening of an *A. annua* cDNA library. Because there are several terpene synthases present in *A. annua* (Bouwmeester et al. 1999), such a PCR may yield a

The cDNA sequence of amorpha-4,11-diene synthase has been submitted to the GeneBank/NCBI data bank with accession number AY006482 and to the European Patent Office under application number 98202854.0 (Wallaart and Bouwmeester 1999)

Abbreviations: DHAA = dihydroartemisinic acid; FDP = farnesyl diphosphate; PCR = polymerase chain reaction

Correspondence to: T. E. Wallaart;

E-mail: mail@genoclip.com; Fax: +31-50-3176790



**Fig. 1.** Proposed biosynthetic pathway of artemisinin in *Artemisia annua*, starting with the amorpha-4,11-diene synthase-catalysed conversion of FDP into amorpha-4,11-diene. The conversion of dihydroartemisinic acid into artemisinin consists of non-enzymatically catalysed reactions starting with the photooxidation (singlet oxygen,  $^1\text{O}_2$ ) of dihydroartemisinic acid, which proceeds through a classical ene reaction with rearrangement of the endocyclic double bond, yielding dihydroartemisinic acid hydroperoxide. Ring cleavage, induced by air oxidation (triplet oxygen,  $^3\text{O}_2$ ), and re-formation then complete the ring structure with artemisinin as the end-product. IPP Isopentenyl pyrophosphate (=isopentenyl diphosphate)

variety of terpene synthase representing PCR products with the same length. This problem can be circumvented by specific enhancing of the expression level of amorpha-4,11-diene synthase. Recently, we reported that dihydroartemisinic acid (DHAA; Fig. 1), a biosynthetic derivative of amorpha-4,11-diene, acts as a scavenger of singlet oxygen ( $^1\text{O}_2$ ; Wallaart et al. 1999a, b). The DHAA, and consequently also its precursor amorpha-4,11-diene, are produced in response to the release of reactive oxygen (Wallaart et al. 1999b). Thus, exposure of *A. annua* to stress conditions in which reactive oxygen species (e.g.  $^1\text{O}_2$ ) are produced, will preferentially induce the transcription of the gene encoding amorpha-4,11-diene synthase (Knox and Dodge 1985; Wallaart et al. 1999b). Here we describe the isolation of a cDNA encoding amorpha-4,11-diene synthase and its functional expression in *Escherichia coli* and *N. tabacum*.

## Materials and methods

### Plant material and growth conditions

*Artemisia annua* L. plants originating from Hanoi, Vietnam (taxonomically verified specimen is deposited at our company; deposit no. Aa002) were grown under sterile conditions on Murashige and Skoog medium (Duchefa, Haarlem, The Netherlands) supplemented with 3% (w/v) sucrose, pH 5.9 and solidified with 0.8% (w/v) micro agar (Duchefa), at 23 °C, >95% moisture, under a 16 h light (ca. 3,000 lx) and 8 h dark cycle. Tobacco (*Nicotiana tabacum* L., cv. Petite Havana SR1) plants were grown on the same medium and under the same conditions (Malaga et al. 1973).

### Total RNA isolation and cDNA synthesis

*Artemisia annua* plants were stressed by placing them at approximately 30% relative humidity (drought stress) and 6,000 lx (photo

stress) for 30 min. After a recovery period of 1 h, total RNA was isolated from young leaves (shoots) by using a DNeasy RNA isolation kit according to the manufacturer's instructions (Qiagen). Impurities (such as chromosomal DNA) were removed from the RNA by incubation with DNase I. (RNase free; Gibco BRL). Copy DNA was synthesised using SuperScript II reverse transcriptase (Gibco BRL), according to the manufacturer's instructions.

### Polymerase chain reaction-based probe generation

Based on comparison of the deduced amino acid sequences of spearmint [(–)-limonene synthase], tobacco (5-*epi*-aristolochene synthase) and castor bean (cadinene synthase) two highly conserved domains were selected (Back et al. 1994) for which a set of consensus inosine-containing, degenerated primers (primers A and B) was synthesised. Primer A was designed based on the highly conserved domain of the proposed active site of the terpene synthases with the characteristic DDXXD sequence motif, in detail: DD(T/I)(I/Y/F)D(A/V)Y(A/G); primer B was designed to the highly conserved domain (D/N)(S/E)(D/E/N)G(K/E)FKE. The sequences of the anti-sense primer A and the sense primer B were respectively: 5'-C RTA IGC RTC RAA IGT RTC RTC-3' and 5'-GAY GAR AAY GGI AAR TTY AAR GA-3'. The letters R (A or G), and Y (C or T) designate IUB codes for variable nucleotide sites, and I denotes inosine.

The PCR was performed using 0.5  $\mu\text{M}$  of each primer, 0.25  $\mu\text{g}$  of template cDNA, and SuperTaq DNA polymerase (HT Biotechnology, Cambridge, UK) according to the manufacturer's instructions. Amplification was performed in a thermal cycler (Mastercycler personal; Eppendorf) as follows: 2 min at 94 °C; 40 cycles of 1 min at 94 °C, 2 min at 40 °C, and 1 min 15 s at 72 °C, followed by 10 min at 72 °C. The amplification products of the expected size (ca. 0.5 kb) were made blunt by the DNA polymerase I large (Klenow) fragment, ligated into *Sma*I-digested plasmid pGEM-7Zf(+) (Promega) and used to transform *E. coli* DH5 $\alpha$  (Gibco BRL).

### Cloning of cDNA

The cDNA was made double stranded by using the RiboClone cDNA synthesis system (Promega). After ligation with *Eco*RI (*Not* I) adapters (Gibco BRL) the double-stranded DNA was ligated into a  $\lambda$  ExCell *Eco*RI/CIAP vector (Amersham-Pharmacia Biotech). For packaging and plating of the cDNA library, the Ready-To-Go Lambda Packaging Kit (Amersham-Pharmacia Biotech) was used. For library screening, 200 ng of the PCR-amplified probe was gel-purified, randomly labelled with  $\alpha$ -[ $^{32}\text{P}$ ]dCTP, according to the manufacturer's recommendation (Random Primed DNA Labeling Kit; Boehringer Mannheim) and used to screen the cDNA library according to Sambrook et al. (1989). The positive clones were sequenced with an automatic LICOR Biotechnology sequencer. Both isoelectric point and molecular weight were calculated from the deduced amino acid sequence using the ExPASy ftp server.

### Expression of the amorpha-4,11-diene synthase cDNA in *E. coli*

The full-length open reading frame of amorpha-4,11-diene synthase, amplified by PCR (under conditions as described above) using primers with *Nco*I and *Bam*HI sites at the ends [forward primer C (5'-GT CGA CAA ACC ATG GCA CTT ACA GAA G-3'), reverse primer D (5'-GGATGGATCC\_TCA TAT ACT CAT AGG ATA AAC G-3')], was inserted into a *Nco*I/*Bam*HI-digested pET 11d expression vector (Stratagene). For expression, this gene construct was used to transform *E. coli* BL21(DE3) (Stratagene). Clones were grown overnight on Luria broth (LB; Duchefa,

Haarlem, The Netherlands) agar plates supplemented with 25 µg/ml ampicillin at 37 °C. Cultures of 50 ml LB medium, supplemented with 100 µg/ml ampicillin and 0.25 mM isopropyl β-D-thiogalactopyranoside (IPTG), were inoculated with these overnight cultures to  $A_{600} = 0.5$  and grown for 3 h at 27 °C. The cells were harvested by centrifugation and re-suspended in 2 ml buffer containing 15 mM 3-[N-morpholino]-2-hydroxypropanesulfonic acid (Mopso) (pH 7.0), 10% (v/v) glycerol, 1 mM sodium ascorbate, 10 mM MgCl<sub>2</sub> and 2 mM dithiothreitol (DTT) (Bouwmeester et al. 1999). An aliquot of 1 ml re-suspended cells was sonicated on ice four times for 5 s with interruptions of 30 s. The sonicate was centrifuged for 5 min at 4 °C in a micro centrifuge (13,000 rpm) and the supernatant was used to assay cyclase activity using 20 µM [<sup>3</sup>H]FDP (1.85 GBq mmol<sup>-1</sup>) as substrate and to carry out SDS-PAGE gel electrophoresis followed by staining with Coomassie brilliant blue (Bouwmeester et al. 1999; Sambrook et al. 1989). Labelled pentane-soluble products were analysed using radio-GC and GC-MS, essentially as described by Bouwmeester et al. (1999).

### Plant transformation

To clone the amorpha-4,11-diene synthase cDNA into a *Bam*HI site of a CIAP dephosphorylated pLV399 plant expression cassette, *Bam*HI restriction sites were introduced by PCR, under conditions as described above, with forward primer E (5'-GAG GAT CCA TGT CAC TTA CAG AA-3') and reverse primer F (5'-ATG GAT CCT CAT ATA CTC ATA GGA-3') (McCormac et al. 1997). The orientation of the insert in pLV399 was checked by double digestion with *Nde*I and *Pst*I. Partial *Kpn*I digestion of this construct yielded a 2.67-kb fragment (containing the amorpha-4,11-diene synthase cDNA in the plant expression cassette) which was subsequently subcloned in a *Kpn*I-cut pCGN1548 binary vector. This construct was subsequently used to transform leaf discs of *N. tabacum* from which kanamycin-resistant plants were regenerated (Ditta et al. 1980; Bevan 1984; Rogers et al. 1986).

### Analysis of transformed plants for amorpha-4,11-diene

After 45 d under climate-room conditions, tobacco plants were collected for analysis of amorpha-4,11-diene. For each sample approximately 2.5 g of leaves was homogenised using a mortar and pestle in 5 ml pentane (p.a.) to which ca. 1 g anhydrous Na<sub>2</sub>SO<sub>4</sub> was added. After grinding, the pentane phase was transferred to a centrifuge tube, and decolorised with activated carbon. After centrifugation for 5 min at 3,000g, the pentane extract was filtered over a short column (<0.5 cm) containing Al<sub>2</sub>O<sub>3</sub> (grade III) and a spatula tip of anhydrous Na<sub>2</sub>SO<sub>4</sub>. The samples were carefully concentrated under a stream of N<sub>2</sub> and analysed using GC-MS, as described by Bouwmeester et al. (1999), in both the scanning mode and the selected ion monitoring mode: *m/z* 119, 121, 189. Concentrations of amorphadiene as measured in the selected ion monitoring mode were calculated by comparing peak areas with the peak areas of an amorphadiene standard measured in both scanning and selected ion monitoring mode, and a *trans*-caryophyllene standard of known concentration, measured in scanning mode.

## Results

### Amplification of a probe by PCR

To obtain probes for screening an *A. annua* cDNA library, a PCR with primers A and B, homologous to conserved sequences of terpene synthases, was performed on cDNA synthesised with RNA obtained from

a drought- and photo-stress-exposed *A. annua*. Sequencing of sub-cloned PCR products, obtained from 12 individual PCRs, revealed that they were all 538 bp long and 100% identical (Fig. 2, underlined part). Computer-assisted comparison (BLAST comparison) showed that this sequence encodes a part of a terpene synthase. On the other hand, the PCRs with cDNA synthesised using RNA isolated from non-stressed plants did not yield any PCR product visible on agarose gels.

### Cloning of the amorpha-4,11-diene synthase cDNA

The 538-bp PCR fragment was subsequently used to screen the *A. annua* cDNA library. The titre of the unamplified library was  $1.2 \times 10^6$  plaque-forming units

	AAATC	5
ATGTCACCTACAGAGAAAAACCTATTCGCCCATCTGCCAACTTCTCCAAAGCATTGG		65
M S L T E E K P I R P I A N F P P S I W		
GGAGATCAGTTTCTCATCTATCAAAGCAAGTAGAGCAAGGGTGGACAGATAGTGAAT		125
G D Q F L I Y Q K Q V E Q G V E Q I V N		
GATTTAAAAAAGAGTGGCGCACTACTAAAGAAAGCTTTGGATATTCCTATGAAACAT		185
D L K K E V R Q L L K E A L D I P M K H		
GCCAAATTTGTTGAAGCTGATTGATGAAATCAACGCCCTTGAATACCGTATCACTTTGAA		245
A N L L K L I D E I Q R L G I P Y H F E		
CGGAGATTGATCATGCATTGCAATGATTTATGAAACATATGGTGATACTGGAATGGT		305
R E I D H A L Q C I Y E T Y G D N W N G		
GACCGCTCTTCCTATGGTTCCTGCTTATGCGAAGCAAGGATATTTATGTTACATGTAT		365
D R S S L W F R L M R K Q G Y V Y T C D		
GTTTTCAATAACTATAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTT		425
V F N N Y K D K N G A F K Q S L A N D V		
GAAGGTGTTGCTTGAGTTGTACGAAGCAACTCTATGAGGGTACCTGGGGAGATTATATTA		485
E G L L E L Y E A T S M R V P G E I I L		
GAAGATGCTCTTGGTTTTACACGATCTCGTCTTAGCATTATGACAAAGATGCTTTTCT		545
E D A L G F T R S R L S I M T K D A F S		
ACAAACCCCGCTCTTTTACCGAAATACAAAGGCGACATAAGCAACCCCTTTGGAAAGG		605
T N P A L F T E I Q R A L K Q P L W K R		
TTGCCAAGATAGAGCGCGCGCAGTACATCTCTTCTATCAACCAAGATTCCTCATAC		665
L P R I E A A Q Y I F F Y Q Q Q D S H N		
AAGACTTTACTTAAACTCTGAAGTTAGAGTTCAATTTGCTTCAGTCACTGCAACAGGAA		725
K T L L K L A K L E F N L F Q S E		
GAGCTCAGCATGTGTGCAATGGTGAAGCTTTCGATATCAAGAGAACGCACTTGT		785
E L S H V C K A K W K A F D I K K N A P C		
TTAAGAGATAGAATTTGTTGAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAG		845
L R D R I V E C Y F W G L G S G Y E P Q		
TATTCGGGCTAGAGTTTCTTCAAAAGCTGTGCTGTATATCTCTTATAGATGAC		905
Y S R A R V F F T K A V A V I T L I D D		
ACTTATGATGCGTATGTTACTTATGAAGAACTTAAGTCTTTACTGAAGCTGTGAAAGG		965
T Y D A Y G T Y E E L K I F T E A V E R		
TGGTCAATTACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATC		1025
W S I T C L D T L P E Y M K P I Y K L F		
ATGGATACATACAGAAATGGAAGAATTTCTTGAAGAGGAGGAAGAACAGATCTATTT		1085
M D T Y T E M E E F L A K E G R T D L F		
AACCTGGGCAAGAAATTTGAAAGAGTTTGTAGAAACCTGATGGTTGAAGCAAAATGG		1145
N C G K E F V K E F V R N L M V E A K W		
GCAATAGAGGACACATCAACCACTGAAGAGCATGATCCAGTTGTAATCATTACTGGC		1205
A N E G H I P T T E E H D P V V I I T G		
GGTGCTAACCTGCTTCAACAACTTGTATTTCTGGCATGAGTGATATTCACAAAGAG		1265
G A N L L T T T C Y L G M S D I F T K E		
TCGTGCAATGGGCTGTCTGCACTCTCTTTTAGATACTCAGGTATACCTGTGCGA		1325
S V E A V S A P P L F R Y S G I L G R		
CGCCTAATGATCTCATGACCCACAGGCGAGCAAGAAAGAAACATAGTTTCATAGAGC		1385
R L N D L M T H K A E Q R K H S S S S		
CTTGAAGTTATGAAGGAATATAATGTCATGAGGAGTATGCCCAACCTTGATTTAC		1445
L E S Y M K E Y N V N E E Y A Q T L I Y		
AAGGAAGTAGAAGATGTGTGAAGATATAAACCGAGAGTACCTCACAACATAAAACATT		1505
K E V E D V W K D I N R E Y L T T K N I		
CCAGGCGCTTATTGATGCTGTGATCTAATTTGTGCGAGTTCTTGAAGTTCAATATGCA		1565
P R P L L M A V I Y L C Q P L E V Q Y A		
GGAAGGATAACTTCACAGTATGGGAGAGCAATACAAACATCTCTAAAGTCCTACTC		1625
G K D N F T R M G D E Y K H L I K S L L		
GTTTATCTCATGATATAGTACCAACTCTCTGTCATAGCCTATCAATATATTTAGAA		1685
V Y P M S I *		
AGGGTTAACTATGACAGTCTCTATGAGAGAAATTTCTCAAGCTATTTGGTGTCTCTGCT		1745
GGCAATTAATAATCAGACGCATAAAATTTGATGAACTATGCGCATAGCTATTTAAAG		1805
TTATATACAACTAAATATTTCAACAATGTGATTATATCTTTTACTTTGTACAAAGCAAA		1865
AGTACACTACTGTTATGTAACATTTTAGTTCTATGATACCTTTAGTTACGAATCGGCTTAT		1925
ATACATTGATACACTTTTATGCAGAAACCOCTAGTAAATAAAAGTCGATATCTGTACT		1985
ACACATATCGCACGAATTTCCGTTTGGCGTTTGTATTTTACGATATGTTATTTAATGAAT		2045
ATGTTTCATGTGGTTGTGCTTAAAAAAA		2076

**Fig. 2.** Nucleotide and deduced amino acid sequences of *A. annua* amorpha-4,11-diene synthase (Wallaart and Bouwmeester 1999). The nucleotide sequence is numbered on the right. The amino acid sequence is given in the single-letter code and is numbered on the left. The sequence of the PCR-generated probe, excluding the oligo-nucleotide sequences, is indicated by underlining

(pfu). Among the 143 initially identified clones, 12 were chosen for further analysis by restriction analysis and nucleotide sequence determination. All 12 clones were found to have an identical nucleotide sequence within their overlapping regions. The nucleotide sequence of the longest cDNA, containing 2,076 nucleotides (GenBank accession number: AY006482), is shown in Fig. 2. The reading frame contains 546 codons, ending with an opal (TGA) stop codon. The calculated molecular mass of the protein was 63.9 kDa, and its calculated isoelectric point 5.66.

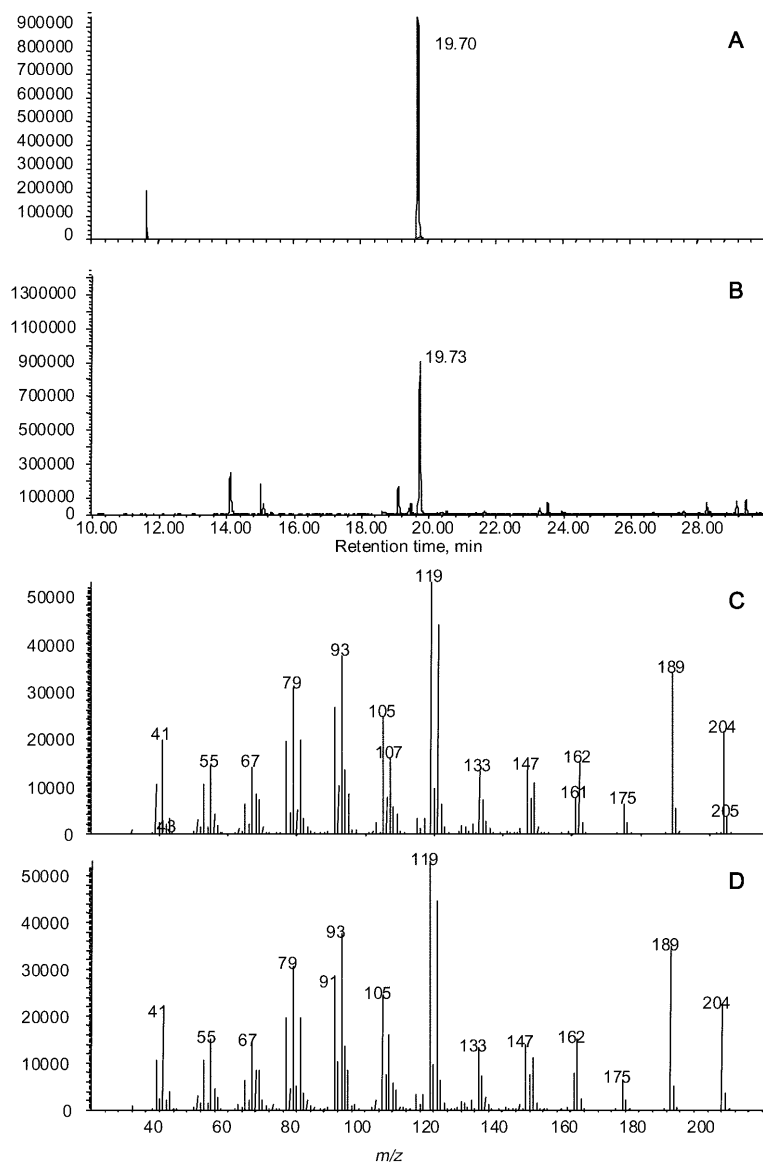
The enzymatic activities of the recombinant protein (as expressed in *E. coli*) were investigated by assays using [ $1\text{-}^3\text{H}$ ]FDP as substrate followed by radio-GC (results not shown) and GC-MS analysis. The GC-MS analysis of the pentane extract indicated the presence of one major (89%) and four minor (together 11%) sesquiterpene products (Fig. 3). The main component, with a retention time of 19.73 min, was in all aspects (retention time and mass spectrum) identical to the reference

amorpha-4,11-diene (Bouwmeester et al. 1999). Of the minor compounds only  $\beta$ -sesquiphellandrene could be identified. No amorpha-4,11-diene was found in the control assays. Therefore we conclude that the cloned cDNA encodes the enzyme amorpha-4,11-diene synthase.

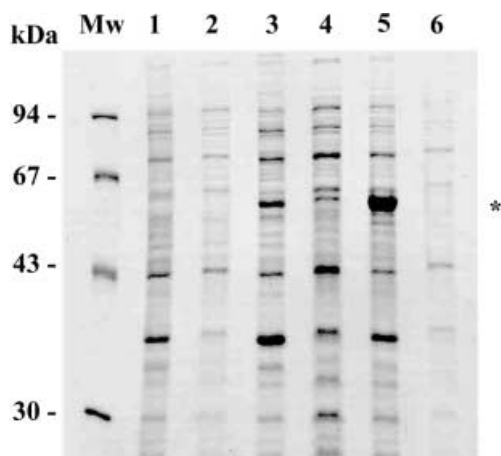
An SDS-PAGE analysis of the total cellular protein of bacterial cells expressing the recombinant amorpha-4,11-diene synthase, revealed the presence of an intense protein band of approximately 50–60 kDa (Fig. 4), which was absent in extracts of the negative control culture. The negative control culture contained the pET 11d vector without insert.

#### *Expression of the amorpha-4,11-diene synthase-encoding gene in N. tabacum*

Transformed tobacco plants appeared normal with respect to morphology and growth. In the homogenates



**Fig. 3A–D.** Analysis by GC-MS of the product (amorpha-4,11-diene) formed from FDP in an assay with recombinant amorpha-4,11-diene synthase expressed in *E. coli* BL21(DE3). **A** Chromatogram of an authentic sample of amorpha-4,11-diene (Bouwmeester et al. 1999); **B** chromatogram of the pentane extract of the cyclase (FDP) assay; **C** mass spectrum of authentic amorpha-4,11-diene; **D** mass spectrum of the major peak in **B**



**Fig. 4.** Analysis by SDS-PAGE of protein extracts from *E. coli* BL21(DE3) expression cells containing the following plasmid constructs: lanes 1 and 2 pET 11d (negative control); lanes 3 and 4 pET 11d containing the tobacco *epi*-aristolochene synthase (positive control). Lanes 5 and 6 pET 11d containing the amorpha-4,11-diene synthase cDNA. Lanes 1, 3, 5 Pellets; lanes 2, 4, 6, supernatants. Mw Low-molecular-weight marker (Amersham-Pharmacia Biotech). The lanes with the pellet fractions of TEAS in pET 11d (positive control) and the amorpha-4,11-diene synthase in pET 11d showed a clear band, indicated with an asterisk, which was not present in the negative control pET 11d

of seven different transformed plants amorpha-4,11-diene was detected at levels of 0.2–1.7 ng per g leaf tissue (fresh weight). Amorpha-4,11-diene was absent in three untransformed control samples.

## Discussion

### Isolation of the amorpha-4,11-diene synthase cDNA

There was no PCR product visible on 0.8% agarose gels when RNA obtained from non-stressed plants was used to synthesise the template. This is consistent with the characteristics of terpene synthases, which are known to occur only in very low intracellular concentrations (Croteau and Cane 1985). The impact of stress on the transcript level of the amorpha-4,11-diene synthase gene was crucial for the success of the PCR. The applied combination of drought and photo stress (inducing elevated levels of  $^1\text{O}_2$  in the plant) turned out to be a highly specific induction of transcription, since only a single terpene synthase sequence, encoding a part of the amorpha-4,11-diene synthase gene, was obtained by PCR. These results support our recent findings in which the biosynthetic derivative of amorpha-4,11-diene, DHAA, is produced in response to the release of reactive oxygen, and acts as an antioxidant by quenching  $^1\text{O}_2$ , yielding artemisinin as a stable end-product (Fig. 2) (Wallaart et al. 1999a, b).

Analysis of the full-length cDNA clone (Fig. 2) around the first ATG codon, at positions 6–8, suggests that this is the initiating codon (Back and Chappell 1995). As is typical of the sesquiterpene synthases of plant origin, amorpha-4,11-diene synthase lacks an

amino-terminal organelle-targeting sequence (Facchini and Chappell 1992; Back et al. 1994; Crock et al. 1997). Therefore, the enzyme is most likely directed at the cytoplasm, which is the site of sesquiterpene biosynthesis, while mono- and diterpenes are synthesised in plastids (Colby et al. 1993). Comparison of the deduced amino acid sequence with sequences of other terpene synthases revealed the aspartate-rich motif (DDXXD) found in most prenyltransferases and terpene synthases, which is thought to be involved in binding the divalent metal ion-chelated substrate (Croteau and Cane 1985; Tarshis et al. 1994). Comparison of the deduced amino acid sequence of the amorpha-4,11-diene synthase cDNA (Fig. 2) with those of other plant sesquiterpene synthases reveals a significant degree of similarity. The sequence shows highest similarity of 67% (50% identity) to a putative sesquiterpene cyclase from *A. annua* (EMBL accession number AJ249561; BLAST according to Altschul et al. 1997). The homology with another terpene synthase from *A. annua*, the monoterpene synthase (3R)-linalool synthase (Jia et al. 1999), showed only 49% (31% identity).

### Functional expression of the amorpha-4,11-diene synthase encoding cDNA

The amorpha-4,11-diene synthase assay with recombinant amorpha-4,11-diene synthase did not yield exclusively, amorpha-4,11-diene, but also some other sesquiterpenes, forming 11% of the total products (Fig. 4). This feature is not unique for recombinant amorpha-4,11-diene synthase. The recombinant terpinolene synthase from grand fir (*Abies grandis*), for example, also yielded more than one product (Back and Chappell 1996; Bohlmann et al. 1999).

The presence of amorpha-4,11-diene in the transgenic tobacco plants demonstrated that amorpha-4,11-diene synthase can be expressed in an active form in a plant. Nevertheless, the concentrations of amorpha-4,11-diene in these transgenic plants were low. These findings are in accordance with those of Hohn and Ohlrogge (1991) who expressed the fungal sesquiterpene synthase 'trichodiene synthase' in tobacco. Also in their experiment, trichodiene could only be detected in trace amounts, ranging from 5 to 10 ng/g FW, which is nearly 10-fold higher than we found. An explanation for these differences may reside in conditional differences under which the plants were grown. Instead of MS medium and climate-room conditions as we used, the plants of Hohn and Ohlrogge (1991) were grown in potting soil under greenhouse conditions. In spite of these differences the sesquiterpenes in both cases are produced in trace amounts. An explanation for these low sesquiterpene levels in general may be the high level of regulation of sesquiterpenoid biosynthesis in tobacco. After treatment with a suitable elicitor, sesquiterpenoids in tobacco increased from undetectable to relatively high levels for a temporary period of time (Facchini and Chappell 1992). Following induction, cell cultures also had increased activity of 3-hydroxy-3-methylglutaryl

coenzyme-A reductase and decreased levels of sterol biosynthesis (Vögeli and Chappell 1988). Elicitor-mediated induction of 5-*epi*-aristolochene synthase has recently been shown to be transcriptionally regulated (Vögeli et al. 1990). The significant alterations in several steps of isoprenoid metabolism after elicitor treatment suggest that expression of a foreign sesquiterpene synthase alone might not be enough for the production of high levels of sesquiterpenoids in tobacco. Sesquiterpenoid levels in *A. annua*, however, are much higher. Specifically the high levels of artemisinin and some of its precursors indicate the presence of a relatively high level of isoprenoid metabolism in this plant (Wallaart et al. 2000). Manipulation of the expression of amorpha-4,11-diene synthase in transgenic *A. annua* should demonstrate whether amorpha-4,11-diene synthase is indeed the main rate-limiting regulatory step in the biosynthesis of artemisinin.

## References

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PST-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402
- Back K, Chappell J (1995) Cloning and bacterial expression of a sesquiterpene cyclase from *Hyoscyamus muticus* and its molecular comparison to related terpene cyclases. *J Biol Chem* 270: 7375–7381
- Back K, Chappell J (1996) Identifying functional domains within terpene cyclases using a domain-swapping strategy. *Proc Natl Acad Sci USA* 93: 6841–6845
- Back K, Yin S, Chappell J (1994) Expression of a plant sesquiterpene cyclase gene in *Escherichia coli*. *Arch Biochem Biophys* 315: 527–532
- Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Res* 12: 8711–8721
- Bohlmann J, Phillips M, Vasanthi R, Sadanobu K, Croteau R (1999) cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (*Abies grandis*). *Arch Biochem Biophys* 368: 232–243
- Bouwmeester HJ, Wallaart TE, Janssen MHA, van Loo B, Jansen BJM, Posthumus MA, Schmidt CO, De Kraker J-W, König AW, Franssen MCR (1999) Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* 52: 843–854
- Colby SM, Alonso WA, Katahira EJ, McGarvey DJ, Croteau R (1993) 4S-Limonene synthase from the oil glands of spearmint (*Mentha spicata*): cDNA isolation, characterization, and the bacterial expression of the catalytically active monoterpene cyclase. *J Biol Chem* 268: 23016–23024
- Crock J, Wildung M, Croteau R (1997) Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha × piperita*, L.) that produces the aphid alarm pheromone (*E*)-beta-farnesene. *Proc Natl Acad Sci USA* 94: 12833–12838
- Croteau R, Cane DE (1985) Monoterpene and sesquiterpene cyclases. *Methods Enzymol* 110: 383–405
- Ditta G, Stanfield S, Corbin D, Helinski DR (1980) Broad host-range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci USA* 77: 7347–7351
- Facchini PJ, Chappell J (1992) Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. *Proc Natl Acad Sci USA* 89: 11088–11092
- Hohn TM, Ohlrogge JB (1991) Expression of a fungal sesquiterpene cyclase gene in transgenic tobacco. *Plant Physiol* 97: 460–462
- Jia JW, Crock J, Lu S, Croteau R, Chen XY (1999) (3R)-Linalool synthase from *Artemisia annua* L.: cDNA isolation, characterization, and wound induction. *Arch Biochem Biophys* 372: 143–149
- Knox JP, Dodge AD (1985) Singlet oxygen and plants. *Phytochemistry* 24: 889–896
- Lütke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J* 6: 43–48
- Malaga P, Sz-Breznovits A, Marton L (1973) Streptomycin-resistant plants from callus culture of haploid tobacco. *Nat New Biol* 224: 29–30
- McCarvey DJ, Croteau R (1995) Terpenoid metabolism. *Plant Cell* 7: 1015–1026
- McCormac AC, Malcolm EC, Fong C-D (1997) pBECKS: a flexible series of binary vectors for *Agrobacterium*-mediated plant transformation. *Mol Biotechnol* 8: 199–213
- Rogers SG, Horsch RB, Fraley RT (1986) Gene transfer in plants: production of transformed plants using Ti plasmid vectors. *Methods Enzymol* 118: 627–640
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York
- Tarshis LC, Yan M, Poulter CD, Sacchettini JC (1994) Crystal structure of recombinant farnesyl diphosphate synthase at 260-Å resolution. *Biochemistry* 33: 10871–10877
- Van Geldre E, Vergauwe A, Van den Eeckhout E (1997) State of the art of the production of the antimalarial compound artemisinin in plants. *Plant Mol Biol* 33: 199–209
- Vögeli U, Chappell J (1988) Induction of sesquiterpene cyclase and suppression of squalene synthase activities in plant cell cultures treated with fungal elicitor. *Plant Physiol* 88: 1291–1296
- Vögeli U, Freeman JW, Chappell J (1990) Purification and characterization of an inducible sesquiterpene cyclase from elicitor-treated tobacco cell suspension cultures. *Plant Physiol* 94: 1860–1866
- Wallaart TE, Bouwmeester HJ (1999) DNA encoding amorpha-4,11-diene synthase. Representative: FHM Van Someren et al. International patent application ID\EP99\06302, European Patent Application No. 98202854.0
- Wallaart TE, van Uden W, Lubberink HGM, Woerdenbag HJ, Pras N, Quax WJ (1999a) Isolation and identification of dihydroartemisinic acid from *Artemisia annua* and its role in the biosynthesis of artemisinin. *J Nat Prod* 62: 430–433
- Wallaart TE, Pras N, Quax WJ (1999b) Isolation and identification of dihydroartemisinic acid hydroperoxide from *Artemisia annua*: a novel biosynthetic precursor of artemisinin. *J Nat Prod* 62: 1160–1162
- Wallaart TE, Pras N, Quax WJ (2000) Seasonal variations of artemisinin and its biosynthetic precursors in tetraploid *Artemisia annua* plants compared with the diploid wild-type. *Planta Med* 66: 57–62
- Webster HK, Lehnert EK (1994) Chemistry of artemisinin: an overview. *Trans R Soc Trop Med Hyg* 88: 27–29